

The Analysis of Water Discharge Frequency for Diyala River D/S Derbendi-Khan Dam

Wisam A. Abidalla¹ Ahmed S. Naser² Ali H. Hommadi³

¹ *Karbala Technical Institute, Al-Furat Al-Awsat Technical University, 56001*

² *Ammara Technical Institute, Southern Technical University*

³ *Alhindyia Project, State Commission For Dam and Reservoir, Iraq*

inm.wsm@atu.edu.iq

Ahmedsaminasset@stu.edu.iq

alihassan197950@gmail.com

Submission date:- 21/7/2019	Acceptance date:- 5/1/2020	Publication date:- 13/1/2020
------------------------------------	-----------------------------------	-------------------------------------

Abstract

The analysis of water discharge frequency for Diyala river D/S Derbendi-Khan Dam have been computed using some statistical models such as " Gumbel, Pearson III, log Pearson III and log-normal III. The models were applied to annual water discharge for Diyala river. The water discharge magnitudes computed for many return periods . Firstly, The models were Compared using different statistical measures " bias, standard error and root mean square error. The Chi-Square test was used to evaluate the goodness of fits for these models. According to this test and using the aforementioned measures, the Log-Normal III distribution could be taken as the best for the water discharge series. This type of model is useful for designing the dam or spillway by computing the magnitude of water discharge flood versus different return periods.

key words: Water discharge frequency, Statistical models, Return period, Flood.

1. Introduction

Sometimes, the hydrology tools were impacted by events, like storms, floods, and droughts. For hydrologic data, the frequency analysis objective is to evaluate the values of extreme events to their occurrence frequency through the using of probability distributions. The results of frequency analysis are useful for several purposes to design bridges and dams.

2. Material and Methods

2.1. Study Area

The study area is located at Diyala River / downstream Derbendi-Khan Dam, in the Sulaymaniyah Governorate of Northern Iraq, Location (Latitude 35° 08' 00" N - Longitude 45° 45' 00" E), as shown in fig. (1). The water discharges in this site were recorded for many years. In this research, the period for recording data of water discharge from (1976 - 2018) [1].

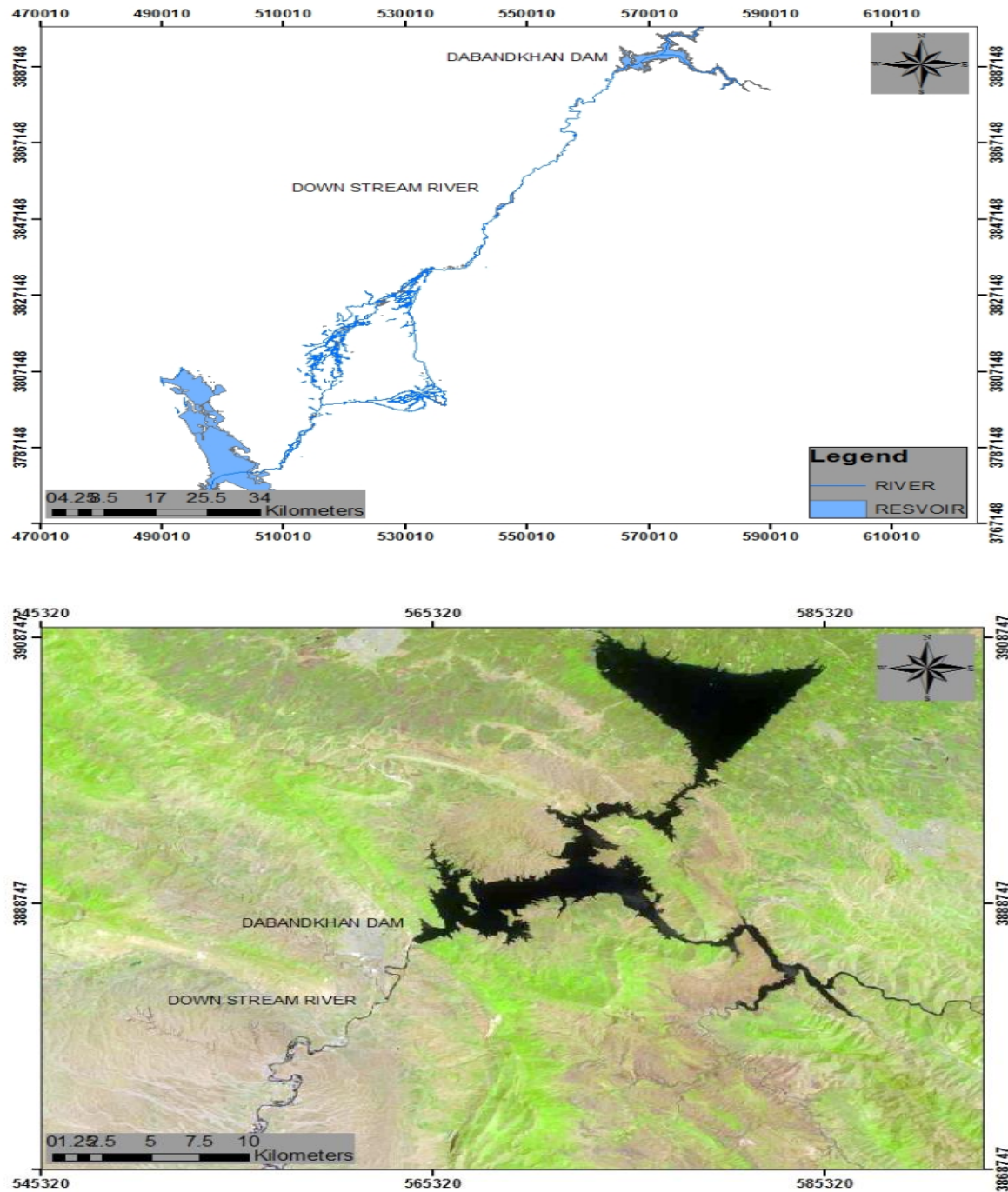


Figure1. Location of Diyala river D/S Derbendi-Khan Dam

2.2. Statistical models

[2] refers that Gumbel distribution can be applied to evaluate maximum events. [3] studied log-normal III. The analytical expression for the standard common probability model was taken by [4] which study the use of moments to evaluate the magnitude of the event and the standard error for different return periods. [5] studied the evaluator on four-wake by the distribution Log Pearson III, Gumbel, Log – Gumbel. [6] studied an estimation of deferent streamflow frequency distributions using the annual peak data. The moment's method used to evaluate the parameters for models. The Gumbel and Log-normal III models are better than the other models. The aim of frequency analysis for the information of hydrology is to evaluate the events

to the frequency of the probability models. The Relationship between period, known- exceedance, $F(x=X)$, with a value of time value can be obtained for 3 parameter distribution as : -

$$P(X \subseteq X) = F(x=x) = \int f(x, \alpha, \beta, \gamma) dx = 1 - \frac{1}{T} \dots \dots \dots (1)$$

where γ , α , β mean location, shape and scale.

$$\mu = \int_1^u .Xf(x).dx = \bar{x} = \frac{\sum x_i}{n} \dots \dots \dots (2)$$

$$\sigma^2 = \int_1^u (x - \mu)^2 f(x) dS = sd^2 = \frac{\sum (xi - \bar{x})^2}{n} \dots \dots \dots (3)$$

$$G = \sum_1^u (X - \mu)^3 f(X) dx / \sigma^3 = c_s = \frac{(\sum (x_i - x))n}{(n-1)(n-2)sd^3} \dots \dots \dots (4)$$

Having selected distribution and estimated its parameters, [7] proposed a general equation to use this distribution in frequency analysis.

$$X_T = \mu + K\sigma \dots \dots \dots (5)$$

where (X_T) is the event magnitude at a given return period, T. (μ) and (σ) are the mean with deviation.

A measure of variability of the resulting event magnitudes is the standard error of estimate. Standard error (S_T), may be written by using method of moments [4] as:

$$S_T = \left(\frac{\partial X}{\partial \alpha} \right)^2 + 2 \left(\frac{\partial X}{\partial \alpha} \right) \left(\frac{\partial X}{\partial \beta} \right) COV(\alpha, \beta) \dots \dots \dots (6)$$

where (α , β , and γ) are the estimated parameters.

2.2.1 Gumbel distribution

The probability for occurrence is the value equal to or less than x when x is variant of the maxima can be given by biggest value [8] such as :

$$p(x \subseteq x) = e^{-e^{-\alpha(x-\beta)}} \dots \dots \dots (7)$$

Parameters β , α known as the location, scale parameters. Probability density function corresponds to equation (7)

$$p(x) = \alpha \cdot e^{(-\alpha(x-\beta) - e^{-\alpha(x-\beta)})} \dots \dots \dots (8)$$

Taking the logarithms value to the base (e) in equation (8) double times, can be found that:

$$x = \frac{1}{\alpha} [\alpha \cdot \beta - \ln(p(x))] \dots \dots \dots (9)$$

Since :

$$Tr(x) = \frac{1}{(1 - p(x))} \dots\dots\dots (10)$$

where $Tr(x)$ is the return period. $p(X)$ is the probability.

$$X = \frac{1}{\alpha} \left[\alpha \cdot \beta - \ln \left(\frac{Tr(x) - 1}{Tr(x)} \right) \right] \dots\dots\dots (11)$$

The parameters α , and β can be found using method of moments [4] as the following equations :

$$\alpha = \frac{1.2825}{\sigma} \dots\dots\dots (12)$$

$$\beta = \mu - 0.45\sigma \dots\dots\dots (13)$$

β , α are scale and location parameters. the μ and σ are mean and standard deviation.

With using the method of moments to evaluate the standard error [4] . The standard error can be calculate by the equation :-

$$S_r = \left[1 + 1.139 K + 1.1 K^2 \frac{\sigma^2}{n} \right]^{\frac{1}{2}} \dots\dots\dots (14)$$

2.2.2 Pearson III distribution

This type of distribution can be written as the following form :

$$pro(x) = \frac{1}{\alpha \Gamma(\beta)} \left[\frac{x - \gamma}{\alpha} \right]^{\beta-1} \exp \left(- \frac{x - \gamma}{\alpha} \right) \dots\dots\dots (15)$$

(α) , (β) and (γ) are parameters to be calculated and $\Gamma(\beta)$ is gamma function .with using of method MoM , the parameter as shown below :

$$\beta = \left(\frac{2}{c_s} \right)^2 \dots\dots\dots (16)$$

$$\alpha = \frac{\sigma}{\sqrt{\beta}} \dots\dots\dots (17)$$

$$\gamma = \mu - \sigma \sqrt{\beta} \dots\dots\dots (18)$$

where β , α , and γ are parameters to be estimated. The frequency factor is given [7] as:

$$k = z + (z^2 - 1) \frac{c_s}{6} + \frac{1}{3} (z^3 - \frac{c_s}{6})^2 - (z^2 - 1) (\frac{c_s}{6})^3 + (z - \frac{c_s}{6})^4 + \frac{1}{3} (\frac{c_s}{6})^5 \dots \dots (19)$$

With Using of MOM method [4] , the standard error is given as:

$$S_r^2 = \frac{\sigma}{n} \cdot \left[\frac{1 + k c_s + \frac{k^2}{2} \left(\frac{3 C_s^2}{4} + 1 \right) + 3K \frac{\partial K}{\partial C_s} (C_s + C_s^3/4 + \dots)}{3 \left(\frac{\partial K}{\partial C_s} \right)^2 \left[2 + 3 C_s^2 + 5 C_s + \frac{5 C_s^4}{8} \right]} \right] \dots \dots (20)$$

$$\frac{\partial K}{\partial C_s} = \frac{Z^2 - 1}{6} + \frac{4(Z^3 - 6Z)}{6^3} C_s - \frac{3(Z^2 - 1)}{6^3} C_s^2 + \frac{4Z}{6^4} C_s^3 - \frac{10}{6^4} C_s^4 \dots \dots (21)$$

where skewness as "Cs" and frequency factor as "K"

2.2.3 Log person III distribution

In distributed a Pearson type III , if the logarithms "lnx" of x variables are variant then the x variable will be distributed as log-Pearson type III [4] .

$$p(x) = \frac{1}{\alpha x \Gamma(\beta)} \left[\frac{\ln x - y}{\alpha} \right]^{\beta-1} \cdot \exp \left[- \frac{\ln x - y}{\alpha} \right] \dots \dots (22)$$

[9] used the Log Pearson distribution III properties to estimate method moments for the active data to give the direct application using relationships as shows in equations :-

$$X = \gamma + \sigma \cdot \beta \dots \dots (23)$$

$$t_y = \sigma \sqrt{\beta} \dots \dots (24)$$

$$z_y = 2 / \sqrt{\beta} \dots \dots (25)$$

where "x" : mean for y = ln x : "ty" : The standard deviation for y = Ln x : "γ, β and α" are coefficients of skew for the event , shape , and scale parameter respectively : "zy" : Is coefficient of the skew for the logarithms.

The T- year event can be evaluated by using Pearson distribution III for sample events from:

$$y_T = \ln x_T = \mu_y + k \cdot \sigma_y \dots \dots (26)$$

where μ_y : mean for y = ln x which equal to \bar{y}

σ_y : Standard deviation for $y = \ln x$: k = The frequency factor .

By using method of moments , the standard error can be evaluated using the same equation in Pearson distribution III to obtain " S_{ty} " in log units from the normal deviate and the coefficient of skew for logarithms of the observed events.

2.2.4 Log normal III distribution

The log normal III distribution of reduced variable $(a - x)$, where a is the lower boundary. The probability of density distribution as shown below :

$$p(x) = \frac{x}{(a-x)\sigma_y\sqrt{\pi}} \exp\left[-\frac{\log(a-x)^2}{\sigma_y}\right] \dots\dots\dots(27)$$

If lower boundary " a " is known , then the reduced variable $(a - x)$ can be used together with the procedures described for the log normal distribution III [10].

With using the Method of Moment (MOM):

$$(C_v)_{x-a} = \frac{\sigma}{x-a} \dots\dots\dots(28)$$

$(C_v)_{X-a}$ which can be calculated by using the following equation :

$$C_s = (C_v)^3_{x-a} + 3(C_v)_{x-a} \dots\dots\dots(29)$$

The parameters σ_y and μ_y can be found by MOM method from : -

$$\sigma_y = [\ln((c_v)^2_{x-a} + 1)]^{0.5} \dots\dots\dots(30)$$

$$\mu_y = \ln\left[\frac{\sigma}{(C_v)_{x-a}}\right] - \frac{1}{2}[\ln(C_v)^2_{x-a} + 1] \dots\dots\dots(31)$$

First expression can be evaluated by using the standard normal deviate as the frequency factor:

$$y_T = \mu_y + z.\sigma_y \dots\dots\dots(32)$$

where " μ_y " , " σ_y " mean and of series in $(x - a)$ so that event " X_T " is : -

$$X_T = a + e^{\mu_y + z\sigma_y} \dots\dots\dots(33)$$

2.3. Chi-Square test

To check the differences between the computed event magnitudes with the observed, Chi-Square test has been applied.[11] define as the expression form of this test as: -

$$\chi^2 = \sum_{i=1}^k \frac{(Q_o - Q_c)^2}{Q_c} \dots \dots \dots (34)$$

where: " k " number of the class intervals , " Q_o " the observed and " Q_c " the estimated number of observations in the class interval. " χ^2 " is a chi-square distribution with k- degrees.

Statisticians such as [12] and [13] recommended that the classes can be combined when the expected number in a class is less than three (3) , therefore the test was included this modification. Standard bias (BIAS) can be used for comparison between the fitted models , the measures were computed as :

$$SE = \left[\frac{\sum (Q_o - Q_c)^2}{N - M} \right]^{0.5} \dots \dots \dots (35)$$

$$RMSE = \sum \left[\frac{Q_o - Q_c}{Q_o} \right]^2 \dots \dots \dots (36)$$

$$BIAS = \sum \left[\frac{Q_o - Q_c}{Q_o} \right] \dots \dots \dots (37)$$

Where : N = Size sample : M = Number of the parameter distribution .

Q_o = The observed discharge : Q_c = The computed discharge:

3. Field data

Values of Max. water discharges for Diyala river D/S Derbendi-Khan Dam were set up for a period of (43) years (1976 – 2018) as shown in Table (1) [1]

**Table (1) Maximum water discharge (cumecs) for
Diyala river D/S Derbendi-Khan Dam , for (43) years [1].**

Year	Discharge (Max.) (m ³ / sec)	Year	Discharge (Max.) (m ³ / sec)
1976	397	1998	660
1977	230	1999	132
1978	313	2000	94
1979	300	2001	71
1980	355	2002	99
1981	342	2003	173
1982	346	2004	170
1983	259	2005	275
1984	172	2006	248
1985	482	2007	184
1986	170	2008	112
1987	289	2009	76
1988	436	2010	217
1989	277	2011	133

1990	187	2012	169
1991	239	2013	157
1992	778	2014	115
1993	231	2015	92
1994	362	2016	245
1995	440	2017	133
1996	183	2018	200
1997	154		

4. Results and Discussion

4.1. Predicted flood magnitudes

Some of the statistical models such as Gumbel, Pearson III, log Pearson III, and Log-Normal III, were applied to evaluate the magnitude of water flood discharge for different return periods. The values of distributions were estimated. These values were calculated by moment, as shown in table (2). Also the magnitude of floods for different return periods and the upper limit, lower limit were estimated. The magnitude of (chi-square) with the decision according to (test) whether the model results were accepted or not. The results are shown in table (3). In the table (2) the parameters indicate which distribution should be accepted. Skewness is more than zero and Kurtosis is calculated from the equation $CK = 3 + 1.5(CS)^2 = 7.38$ [4] is close to the Kurtosis computed from the data table (1) , so Log-Normal III distribution could be accepted as the best. The skewness of the logarithm of data should be more than zero for log Pearson III distribution [4] so it also could be accepted as the best. As for Gumbel distribution, the skewness and Kurtosis should be 1.14 and 5.4 respectively [4] so this distribution could not be regarded as the best.

Table (2 - 1) Estimation of Parameter for peak water discharge data

Mean (\bar{x}) (cumecs)	Standard deviation sd (cumecs)	Skewness Cs	Kurtosis Ck
248.81	146.50	1.71	6.60

Table (2 - 2) Estimation of Parameter for peak water discharge of logarithm data

Mean (\bar{y}) (cumecs)	Standard deviation sdy (cumecs)	Skewness Cs	Kurtosis Ck
5.37	0.55	0.11	2.94

Tables (3 - 1) Return periods versus water discharge magnitudes, standard error, upper - lower limit / Type Gumbel Distribution

Return Periods (years)	Water discharge Magnitudes (m³ / sec)	Standard Error (m³ / sec)	Upper Limit (m³ / sec)	Lower Limit (m³ / sec)
1	128.532	39.31	167.84	89.22
2	224.755	40.19	264.95	184.56
5	354.220	67.69	421.91	286.53
10	439.937	91.42	531.36	348.52

20	522.159	115.49	637.65	406.67
50	628.587	147.50	776.08	481.09
100	708.34	171.82	880.16	536.52

Tables (3 - 2) Return periods versus water discharge magnitudes, standard error, , upper - lower limit / Type : Pearson III Distribution

Return Periods (years)	Water discharge Magnitudes (m ³ / sec)	Standard Error (m ³ / sec)	Upper Limit (m ³ / sec)	Lower Limit (m ³ / sec)
1	131.147	110.15	241.30	20.99
2	210.353	71.23	281.58	139.12
5	343.862	57.30	401.16	286.56
10	440.141	122.44	562.59	317.70
20	535.149	185.54	720.69	349.61
50	660.594	260.72	921.32	399.87
100	755.950	311.68	1067.63	444.27

Tables (3 - 3) Return periods versus water discharge magnitudes, standard error, , upper - lower limit / Type Log Pearson III Distribution

Return Periods (years)	Water discharge Magnitudes (m ³ / sec)	Standard Error (m ³ / sec)	Upper Limit (m ³ / sec)	Lower Limit (m ³ / sec)
1	134.708	27.07	161.77	107.64
2	211.900	37.68	249.58	174.22
5	337.854	62.30	400.16	275.55
10	433.694	97.04	530.74	336.65
20	534.493	151.87	686.36	382.62
50	678.581	259.82	938.40	418.76
100	797.303	372.26	1169.56	425.05

Tables (3 - 4) Return periods versus water discharge magnitudes, standard error, upper - lower limit / Type : Log Normal Distribution

Return Periods (years)	Water discharge Magnitudes (m ³ / sec)	Standard Error (m ³ / sec)	Upper Limit (m ³ / sec)	Lower Limit (m ³ / sec)
1	132.652	0.17	157.21	111.93
2	217.142	0.15	251.27	187.65
5	344.424	0.17	408.19	290.62
10	435.010	0.20	529.73	357.23
20	525.877	0.22	657.83	420.39
50	649.543	0.26	840.20	502.15
100	746.955	0.28	989.34	563.96

When we use measures such standard error (Se), root mean square error (Rmes) and (BIAS), smallest values for these measures lead to the best fit. From Table (4) the Log Normal III distribution .

Table (4) the Standard error , the mean square and the biass for four models using max. discharge data.

Type	SE	RMSE	BIAS
Gumbel I	0.6904	0.3492	1.2096
Pearson III	0.7348	0.7120	1.4684
Log Pearson III	0.4806	0.2727	1.1127
Log normal III	0.4202	0.1848	1.0615

4. 2. Goodness of Fits

Mean and standard deviation are used to describe a set of data or observations. These statistics are estimated from samples. Sometimes the samples may be unrepresentative and may, therefore, lead to estimates that are too high or too low.

This estimation will be of no use if they differ from expected values by more than certain prescribed limits. It is, therefore, necessary to test the statistics to see whether their difference is significant or not. Such tests are called the tests of significance.

The more important one is χ^2 – test .The χ^2 – test can be carried out by using the numerical integration to find the value of P (Probability of deviation) if this probability is equal to or less than a given probability value, then that the deviation is significant at the given probability level.

The computed value of χ^2 is used to determine the probability that the deviation would be larger than or equal to computed value. This can be calculated by using the χ^2 distribution as follows:

$$P = P_r(\chi^2 \geq C) = \int_0^\infty f(\chi^2) dx = 1 - \int_0^C f(\chi^2) dx^2$$

$$\text{Where } f(\chi^2) = \frac{1}{2^{d/2} \Gamma\left(\frac{d}{2}\right)} \left[(\chi^2)^{\left(\frac{d-2}{2}\right)} \right] e^{-\chi^2/2}$$

$f(\chi^2)$: is the probability density function of χ^2 : d = degree of freedom: The result given in table (5) shows that Log Pearson III is better model then the other distributions [14] .

Table (5) Chi-square goodness of fit (χ^2) for peak flood data

Type	(χ^2) Chi -Square computed	(d) degree of freedom	Decision
Gumbel I	3.98	4	Accepted
Pearson III	3.49	3	Accepted
Log Pearson III	4.21	3	Accepted
Log Normal III	1.30	2	Accepted

5. Conclusions

With using the results by comparison of fits for measures " SE, RMSE , BIAS and the Goodness of fit " in the analysis for water discharge frequency, and with these methods, the Log-Normal III distribution could be regarded as the best model for the water discharge data for Diyala river D / S Derbendi-Khan Dam.

Conflicts of Interest

The author declares that they have no conflicts of interest.

References

- [1] Discharge for selected gauging stations in Iraq, " The State Commission for Dams and reservoirs " hydrology section , Baghdad – Iraq , 2018.
- [2] Trodorvic,k.and Rorusele, " the Some Problems for Flood Analysis "Wat. Res. Vol (8), No.(6) , 1971.
- [3] Borgs,S.J.,Lettenmair, Bates,C.L, " the Properties of The 3 , the Log-normal Distribution "Wat. Res. , Vol.(3),No.(1), 1978.

- [4] Kite, W.G., "The Flood and Risk in Hydrologic" The Water resource, Colorado, 1977.
- [5] Haktanir, T. " Comparison of Various Flood Frequency Distribution Using Annual Flood Peaks Data of River in Anatolia", Journal of hydrology , Vol. (136), No. (1) , 1991.
- [6] Krczeera, R, " Roubst Flood Frequency " Wat. Res. Res. Vol. (22), No.(1), pp. 317-334, 1982.
- [7] Haktanir, " The Evaluation of Distributions for Flood F.", hydrologic Journal, Vol. (38), No. (1) , 1993.
- [8] Chow, V.T., Maidment, D.R., and Mays, L.W, "Applied Hydrology " McGraw – Hill, New York., 1988.
- [9] Bobee, B.," The Log-Pearson type III distribution and its Application in Hydrology " Wat.Res.Res.,Vol.(11), 1975.
- [10] Pilon, J.P, with Aed mowesk, K. "Asymptetic Veriance of F. Qintil Log- Pearson 3 with Historical " Journal. for hydrologic 155, PP. 58-503, 2016.
- [11] Levin, R.I.,and Rubin, D.S., "Statistics for Management" Wesle Publishing Company, 1994.
- [12] Yavjavich, R "The Probability,Statistical in Hydrologic" water publications Colorado, U.S.A, 1999.
- [13] Haan, T.C., " The Statistic Theory in Hydrologic", Sewo stat university prees., 1977.
- [14] Corder,G.W.,Foreman,D.I,"Nonparametric Statistics for Non- Statisticians,A Step-by-Step Approach",Wiley , 2009.

تحليل تردد التصاريف المائية لنهر ديالى عند مؤخر سد دربندخان

وسام عبد العباس عبد الله أحمد سامي ناصر علي حسن حمادي

1- معهد كربلاء التقني، جامعة الفرات الأوسط التقنية، 56001

2 -معهد العمارة التقني، الجامعة التقنية الجنوبية.

3 -مشروع الهندية، الهيئة العامة للسدود والخزان، العراق

alihassan197950@gmail.com Ahmedsaminasset@stu.edu.iq inm.wsm@atu.edu.iq

الخلاصة

تم تحليل تردد التصاريف المائية لنهر ديالى مؤخر سد دربندخان، باستخدام عدد من التوزيعات الإحصائية مثل توزيع كامبل ، توزيع بيرسن الثالث ، توزيع بيرسن اللوغاريتمي الثالث ، التوزيع الطبيعي اللوغاريتمي الثالث لإيجاد قيم التصاريف المائية الفيضانية ولفترات عودة مختلفة ، هذه التوزيعات قورنت أولاً فيما بينها باستخدام مقاييس إد صائية مثل مقياس الانحدار، الخطأ القياسي و جذر معدل مربع الخطأ. بعد ذلك تم تقييم دقة هذه التوزيعات باستخدام اختبار مربع كاي . وطبقاً لهذا الاختبار يتبين أن توزيع الطبيعي اللوغاريتمي الثالث يمكن أن يعد أفضل توزيع بالنسبة للتصاريف الفيضانية لنهر ديالى . وهذا النموذج ضروري لإيجاد العلاقة بين التصاريف الفيضانية وفترات العودة لتصميم السدود أو المطافح المائية .

الكلمات الدالة: - تردد التصاريف المائية، النماذج الإحصائية، فترة العودة، الفيضان.